Appendix C
October 1992 DemoGraFX Interoperability Evaluation to ACATS

FCC WP4
Review Board Member Interoperability Evaluation

Gary Demos

6 October 1992

Executive Summary

The following discussion is my evaluation of the advanced television (HDTV) proposals before the FCC as part of the WP4 Interoperability review process.

To summarize my comments:

* I asked a number of questions of the proponents, many of these questions were ignored or answered negatively

* The FCC testing process, which is nearing completion, was flawed in that many crucial issues of interoperability were not tested. This is partly due to the fact that the testing process was begun when all the systems were analog. The testing process should be redesigned in light of digital technology and interoperability issues.

The most critical interoperability problems are:

* Two of the digital systems and the one analog system are interlaced, with non square pixels. These systems should be rejected.

* All of the systems are at 59.94 Hz or 60.0 Hz. This is not compatible with computer display interoperability, requiring greater than 70 Hz. No system should be accepted at 59.94 or 60.0 Hz.

* The apparent ability to send movies (24 fps) at higher resolution has not been explored

* The header proposals are not universal, but are buried within the HDTV packet formats

* The packet structures are untested with ATM or other packet networks

* System modularity, which would allow the best of each system, has not been evaluated.

* None of the systems is scalable downward. A lower resolution subset would allow low cost reception at resolution below full quality, but potentially better quality than NTSC.

* Still imagery communication has not been tested, although apparently feasible.

* It could be significant if a partial screen update capability were provided. Some of the systems may have this capability, but it has not been explored or tested. This would allow reduced bit rate presentation of high resolution images during times when other data is being sent (like address authorization lists), or for use by lower data rate devices.

* Compression quality is rapidly advancing. We need only examine the recent two years to get a feel for the rapid pace of development. No proponents offered proposals for handling major advancements in compression technology. In ten years, it is likely that all of the proposed systems will be obsolete in their approach. How will we handle the extensibility issue in order to allow graceful adoption of future advancements in image compression and representation?

* The FCC process has encouraged innovation up to this point by stimulating digital systems to be proposed by the proponents. However, at this point what is required is cooperation and system modification in an orderly manner, with subsequent testing of any such modifications. The FCC process as now constituted <u>IS DESIGNED TO PRECLUDE</u> such cooperation, system modification, and subsequent testing! Thus, the work of this interoperability review board is likely to have little significance unless the FCC process is amended to take interoperability seriously.

Summary Of My Evaluation Of The Proponents

My evaluation is presented in much more detail below. However, a summary of my view of each system is presented here:

		Strengths	*Weaknesses*
Accept If Modified:			
	AT&T/Zenith	Progressive Scan	59.94 Hz
		Square Pixels	No 24 Hz mode
		Ability To Correct ATM	Insufficient ATM Correction
		Prioritized Data	Odd Format Size (523/262)
		8 x 8 small motion vector blocks	Regional leak not fully tested
		Sub-Pixel motion resolution	Buried Header Not Universal
		Vector Quantization	Partial Image Update Not Tested
		Easy NTSC DownConversion	Dependent Audio Channels
		Digital	Not Scalable
	MIT/ATVA	Progressive Scan	59.94 Hz
		Square Pixels	No ATM Mapping
		24 Hz Film Mode	Data Not Prioritized
		Independent Audio Channels	Odd Format Size (525 Datalines)
		Sub Pixel motion resolution	No Header Proposed
		Both 8 x 8 and 16 x 16 motion	Partial Image Update Not Tested
		Easy NTSC DownConversion	Sliding Panels
		Digital	Not Scalable
Reject As Not Interoperable:			
	ATRC	Packet Structure	Interlaced Scan
		Prioritized Data	Non-Square Pixels
		Digital	Provision For Square Confused
			59.94 Hz
			Header Under ATRC Packets
			Dependent Audio
			16 x 16 coarse motion only
			Not Scalable
	G.I.	Digital	Interlaced Scan
			Non-Square Pixels
			59.94 Hz
			No Header
			Dependent Audio
			16 x 32 coarse motion only
			Sliding Panels
			Not Scalable
	NHK		Analog
			Interlaced
			59.94 Hz
			No Header
			No Data Formats
			Not Scalable

Unanswered Questions

* I submitted a list of questions which I was concerned with to the proponents prior to the review meeting. A copy of my questions are attached to this evaluation. My questions were answered in writing and by

presentation by MIT. None of the other proponents answered the questions directly, although some of the questions were answered in the course of presentations.

* In general, the answers which I was able to get fell into the following categories:

1) Seems reasonable, have potential solution, but haven't tried it

a) Could be tested without undue difficulty

b) Would be difficult to test with existing prototype equipment

2) Good issue, haven't considered it.

3) Not able to support the item concerned.

* Some of the proposals for solutions seem too preliminary to be plausible

I will cover these issues in more detail here.

Testing Process

* There have been substantial flaws in the testing process from an interoperability point of view. These are:

* The ATTC only provided 59.94 Hz signals to the proponents. No other rates were explored.

* 24 Hz signals were not supplied even though some of the systems can accept and transmit at 24 Hz

* Increased resolution at 24 Hz vs higher rates has been proposed, but has not been tested.

* 70+ Hz computer display compatibility was not tested in any way

* Still frames were not tested. Quality, how many per second, etc. should be tested.

* No flexibility in performance was tested. Such flexibility would include:

*exploring the limits of pixels per second (Jae Lim's comments were that he could go as high as 100 MPixels/second).

* exploring conditional replenishment

* exploring alternate frame rates

* Potential modularity of systems was not tested. Systems were tested only as a whole.

* Header mechanism has not been tested for reliability (error performance), universality, usefulness.

* ATM mapping and error performance was not tested.

* Computer text and graphics on screen was not tested (PC or computer workstation display). Overlay plane implementations (if used) were not implemented or tested.

* Scalability not proposed and not tested.

* Extensibility not tested. Only proposed verbally.

* NTSC down conversion not tested for quality.

* "Production" or "Contribution" quality extensible superset not tested, although proposed verbally.

* The "high priority subsets" used in ATRC and AT&T/Zenith were not tested for viewability by themselves, although private demonstrations of viewability have been promised.

* The distribution of an address list for pay-per-view cable use would require a lower bit rate to be used for still image or partial screen update during the distribution. The distribution of a million subscriber address list should be tested, and the reliability of delivery should be measured in terms of the number of subscriber errors. The quality of the image during the distribution should be tested.

* Encryption may affect the error performance. The impact of the use of encryption on the quality of the picture should be tested. The protection afforded by proposed encryption schemes should be evaluated. The interaction of headers and encryption should be tested, including "in the clear" headers in an otherwise encrypted data stream for such uses as authorization of decryption to new users being added on.

* A key use of advanced television may be professional consultation and collaborative work. A multi person teleconference should be tested. An example might be a medical consultation.

* Some of the system proponents indicated compatibility with digital compressed NTSC proposals to Cable Labs. Such compatibility should be tested for receiver compatibility, distribution compatibility, and interoperability compatibility. Some claimed that the compressed NTSC standard would only be selected after the advanced television standard. This should be verified if true, but testing for interoperability and compatibility will still need to occur prior to selection of any given compression algorithm.

* The use of a VCR, laser disk, or other devices implies a digital port, wire, and signal. The specifications, reliability, and efficiency of such connections should be evaluated prior to selection of any such signal designations. If industry standard digital formats are used (such as IEEE P1394), then these signals should be tested on these formats.

* There was substantial discussion of higher quality signal feeds. Channels which might have wider bandwidth than 20 Mbits per second include ATM networks, long haul fiber (e.g. SONET), cable TV, satellite transponder channels, cellular digital, OFDM, or other networks. The proposals should be tested with these other data rates for compatibility and quality. The mapping of reduced data rate extractions of 20 Mbits/sec from higher data rates should be tested. The methods for simultaneously generating 20 Mbits/sec in addition to higher rates from a source signal should be tested.

* Some people, such as Jim Clark, Chairman of Silicon Graphics, feel that 3-D graphics will be routinely affordable for home receivers and computers. No testing of 3-D graphics interoperability was investigated.

Test Facility

The ATTC was not designed for much of the above testing.

There are systems in existence which could form the basis of a test facility which could be used to test the above items.

In particular, format independent disk arrays, frame buffers, displays, projection systems, and signal processors exist which could be utilized.

It would be highly desirable if appropriate testing could be accomplished toward the broader focus of interoperability.

The testing which has taken place up to this point has been from an entertainment-only point of view.

Category Comments

Interoperability

Progressive scan and square pixels are far preferable.

Problems with Interlace

960 line interlace has problems downcoverting to 480 interlaced NTSC. This format is only useful for upconverting NTSC 480 lines to 960 interlace (with artifacts, using crude replication).

960 interlace cannot be used as basis for augmentation scalability to higher levels. No interlace format can form the basis of an augmentation pyramid. (see more complete discussions accompanying ATRC proposal evaluation, below)

Scalability

None of the systems is scalable (downward).

Scalability is an issue for cost of initial receivers. Comments concerning the cost of scalability ignore the reason for downward scalability, which is to provide lower cost receiving devices.

Scalable bit rates would be useful for:

- peripherals
- disk
- tape
- multi-speed CD-ROM
- busses
- memory bandwidth
- etc.

Most proposals claimed much better quality at 40 to 180 Mbits/second. As bandwidth increases this sounds promising. However, the problems found by the SMPTE study group on compression, which I chair, with compression compatibility suggest that this requires testing.

Some proposals claim to be able to provide variable data rate. Other than AT&T/Zenith, which already operates at variable data rate, none of the other systems has been investigated in a variable data rate operational mode.

Several proponents proposed the 8 x 8 DCT basic DC coefficients for a low resolution picture at 160 x 90 or 180 x 60. Some mention was made of using the first harmonics to achieve double these resolutions. The likely blocking and artifacts which might result from such extractions suggest that testing and verification of the usefulness of these low quality pictures is required.

MIT claimed 2 x 2 or 4 x 4 extracted pixel agregations, using many more DCT coefficients. How bad would this look with ringing, block edges, and other DCT artifacts?

NTSC quality picture in the portion of the signal used as the low-priority signal would be highly desirable. For the progressive scan systems, a progressive scan enhanced NTSC and wide NTSC would naturally result. This was not proposed by any proponent.

NTSC cable labs proposal compatibility was discussed. Some proponents indicuated that they intend to use the same decompression system for compressed NTSC as well as for advanced television.

Computer Compatibility

Some data formats are more flexible than others. The ATRC data format was most flexible, although the AT&T/Zenith format was more robust to ATM packet loss.

Some proponents indicated that their systems where flexible in formats, and that these systems could be adjusted to different scanning parameters and temporal rates.

The temporal rate problems with 59.94 Hz result in significant interoperability problems with computer CRT displays, which must refresh in excess of 70 Hz. The choice of scanning parameters and temporal rates of the HDTV systems pre-dated any all-digital proposals. Since all of the existing proposals have embraced 59.94 Hz since the time when their formats were analog, none of them are interoperable with the display rates required for computer display. This includes both the presentation of computer display information on a receiver which is primarily used for entertainment, news and sports, as well as the display of entertainment, news, and sports on a computer display.

Thus, 59.94 Hz is unnacceptable from an interoperability point of view. The proponents who have indicated that their systems can adapt to 72 or 75 Hz should be allowed to make those modifications and re-test their systems in order to provide appropriate interoperability with computer uses.

Frame buffer packings in some of the systems focused solely on YUV packings. In computer displays, RGB formats are most common. Overlay planes are also used. The memory organizations and packings for such computer styles of displays should be evaluated in addition to the YUV packings which are most commonly used for entertainment, news, and sports.

Specific comments with respect to each system follow below. However, only the AT&T/Zenith and M.I.T./G.I./A.T.V.A. /C.C.D.C. systems are potentially computer interoperable due to square pixels and progressive scanning. This interoperability of AT&T/Zenith and MIT would depend further on 59.94 Hz being abandonded in favor of 70+ Hz and compatible rates.

The ATRC proposal is unacceptable due to a false notion of what interoperability means. Allowing incompatible formats, including interlace and non-square pixel formats, completely prevents interoperability.

Packets and ATM Compatibility

The ATRC packet proposal looks interesting, and is easy to understand. However, one ATM 48 byte packet being dropped takes out a full 148 byte ATRC packet (made up of 3 ATM packets)

AT&T/Zenith can correct a small number of dropped or bad ATM packets. Each packet is augmented by a block-wise reed solomon code. This is potentially much better. However, the AT&T/Zenith slice/packet structure is odd, with 262 and 523 as numbers within the structure (odd numbers derived from ntsc?).

G.I. and M.I.T. did not propose packet structures or mappings onto ATM.

Header

No header proposal provided a Universal Header, as described in the SMPTE task force report. Each proposal required inflexibly decoding the data block structure of that format, and its error correction and sync codes, before finding the header.

A universal identifier should provide universal sync and indentification, and shouldn't be buried within a particular structure.

The proposals as to how the header would identify internal data were unclear, since headers are buried inside format structures. It is necessary to go inside a complex formatted fixed block, find some bits which tell what the block is. This is called the Header in most proposals. It is not a very robust header proposal.

The universal header should allow foreign data types, such as an independent low bit rate audio channel, to share the data bandwidth with the advanced television picture data. Such foreign data should not be required to be packed within the advanced television data packet and error correction structure, which may not be suitable.

Further, the header must be completely reliable in order to perform its function. None of the proposals appeared to provide sufficient error protection to support a header. This will have to be investigated and tested thoroughly before a header mechanism can be deemed to function properly.

Text and Graphics

Square pixels and progressive scan are really required for text and graphics. Those who addressed this issue indicated that they would implement text and graphics using an overlay plane structure. Such a structure would be constructed as RGB or YUV plus overlay. The underlying RGB or YUV would need to be square pixel and progressive scan for compatibility with the square pixel progressive scan text and graphics.

70+ Hz is an absolute requirement for CRT displays of text and graphics. Thus, the proposed rate of 59.94 Hz cannot be used to support text and graphics.

Text and graphics presentation is relatively untested in all systems, especially text and graphics typicall of current window-based computer screens.

The error correction was not sufficient in any proposals for conveying text and graphics in the usual methods of display postscript, quickdraw, windows, X-windows, etc. These formats are extremely error intolerant. Experimentation and testing will be required in this area.

Many proposals seemed to favor data in aux data areas. Some proposals protect these areas with more error correction and more noise immunity, others do not. Error intolerance of most screen-driving formats not adequately addressed.

MIT indicated that they should be able to support numerous fonts? Other proponents made no comment in this regard. The support of numerous fonts, especially for foreign language character sets, in addition to type faces and sizes, could be significant. Such issues should be tested.

3D graphics support was not discussed or enabled by any of the proponents.

The ability to send a transparency channel (alpha) was not proposed or described by any of the systems, even though this might be a key capability for providing data rate flexibility and interactivity.

Still Frames

Several per second (3) seems to be typical. Are there ways to improve this? How good are still frames which are extracted from the moving image stream?

Quality, number per second, photo-CD compatibility, are some of the issues which should be tested.

Interlaced formats are no good for stills.

Extensibility (upward)

The use of an augmentation/residual approach, as proposed by MIT and AT&T Zenith seems like a good plan.

Multiple 6 Mhz augmentation channels seems useful.

This was not tested by any proponent, although extensibility is an important issue.

Latency

There is a disparity between receiver latency and end-to-end latency. Some of the proponents only quoted latency as receiver latency, and others quoted end-to-end latency.

The Zenith/AT&T number of 250 ms may be typical of G.I. and M.I.T., each of which claimed less than100 ms. ATRC has poor latency due to use of mpeg bi-direction coding format (450ms)

VCR

G.I. has demonstrated 2 Hr 8mm Toshiba VCR. AT&T/Zenith claims to have VHS version working in the lab. Trick modes, although discussed, have not be demonstrated, and no mechanism was proposed for how they might be accomplished.

The main issues for VCR interoperability are probably availability of ancillary data, and index directories at the beginning of the tape or using some other mechanism. The availability of multiple independent audio channels could also be helpful for multiple foreign languages, etc. Is there a mechanism to provide an additional closed captioning window? Could alternate views or 3-D stereoscopic viewing be supported within the proposed format?

Computer Standards

Existing standards have not been evaluated for suitability. Examples of potentially suitable formats include IEEE P1596 for baseband, IEEE P1394 for compressed. ANSI X3T9 FCS and HIPPI for baseband in studio

Temporal Rate Issues

8 x 8 motion vector size better with AT&T/Zenith for motion vector future potential. M.I.T. can provide both 8 x 8 and 16 x 16. The usefulness of motion vectors for temporal rate adjustment has not been determined. The proponents are cautious. Thus, we should not count on these motion vectors providing temporal rate flexibility. However, the smaller 8 x 8 blocks provide more possibility that someday such flexibility may be

possible. Unfortunately, the G.I. and ATRC block sizes of 16 x 16 and 32 x 16 are too large to make temporal rate flexibility using motion vectors even be a possibility.

24 Hz compatibility of the data is offered by all of the proponents except AT&T/Zenith. M.I.T. in particular mentioned the possible of significantly increased resolution, mentioning 2560 x 1440 as being possible. Such a highly desirable option as higher resolution for motion pictures should be fully explored and tested. Would even higher resolution be possible at even lower rates for still frames?

M.I.T. ATRC, and G.I. indicated that they could present 24 Hz source imagery for display at 72 Hz. ATTC testing center did not provide or allow source material at 24 Hz. This is a major oversite in the testing process given the crucial nature of 24 Hz motion picture film compatibility. Even 80% of prime time television is shot on 24 Hz film.

The temporal rate problems with 59.94 Hz result in significant interoperability problems with computer CRT displays, which must refresh in excess of 70 Hz. The choice of scanning parameters and temporal rates of the HDTV systems pre-dated any all-digital proposals. Since all of the existing proposals have embraced 59.94 Hz since the time when their formats were analog, none of them are interoperable with the display rates required for computer display. This includes both the presentation of computer display information on a receiver which is primarily used for entertainment, news and sports, as well as the display of entertainment, news, and sports on a computer display.

Thus, 59.94 Hz is unacceptable from an interoperability point of view. The proponents who have indicated that their systems can adapt to 72 or 75 Hz should be allowed to make those modifications and re-test their systems in order to provide appropriate interoperability with computer uses.

Conditional Replenishment

64 x 48 slices in AT&T/Zenith sound promising. Do other proposals have small regions which are identified by packets which can be controlled at independent rates? Could the AT&T/Zenith system or other systems be tested for this feature?

Updating partial screens would be good for lower rate media than 20 mbits. Such a lower data rate would be required for pay-per-view address lists and other auxiliary data uses such as 3-D graphics or computer text and graphics overlay.

MPEG

Why so much discussion of MPEG? MPEG may not be the best compression technology compared to other possibilities. Motion compensated DCT may not ultimately be the best for that matter.

Group of frames, as used in MPEG, is not a good structure for latency, memory usage during decoding, or for decode/manipulate/re-encode cycles.

MPEG is very low quality. 320 x 240 (half in each direction of NTSC, quarter of the number of pixels) in most implementations. Frame rate is 30 Hz progressive. Also has a 24 hz defined mode. No 70+ hz rates defined in MPEG-1.

mpeg-1 may be too poor, inefficient, costly, and asymetric to ever make it in the market place. Mpeg-2 may be much better and may include an mpeg-1 subset.

MPEG-2 may be much more interesting. still in development. unsure where it will be going

Strong focus of MPEG-2 is on supporting interlace. This is misguided.

mpeg very asymmetric. much harder to encode than decode

Approach of AT&T and GI to MPEG to propose their HDTV formats to MPEG sounds somewhat promising.

Strong interest in scalability in MPEG-2

If US took lead with scalable HDTV down through some form of enhanced NTSC (progressive scan, no color subcarrier and wide NTSC), could help lead mpeg 2 or mpeg 3

Interactive Uses

Partial screen update (conditional replenishment) could be important for interactive use with reduced bandwidth.

Some of the systems have hidden conditional replenishment vs explicit conditional replenishment. The conditional replenishment and manipulation of the screen area must be made explicit if interactive and reduced data rate uses are to be supported. Such interoperability might become a significant requirement. Reduced data rate devices, congested networks, interactive uses. addressing lists, and other uses may require partial or slow screen update using reduced bandwidth.

Advanced Television Proposal Comments

AT&T/Zenith

Tiles and Blocks

8 x 8 better for motion vectors for future potential of temporal interpolations. Even smaller subdivisions would be even better, like 2 x 2 or 4 x 4, but 8 x 8 is smallest region motion vector provided for by AT&T/Zenith proposal. The use of motion vectors for temporal rate flexibility is speculative, and may not be feasible for many years. However, the use of smaller blocks is helpful in providing at least the possibility of such temporal flexibility.

64 x 48 slice tile size appears promising for conditional replenishment. However, independent temporal update control of individual slice tiles has not been tested.

Data Packets

Block Reed Solomon codes, as proposed, will correct packet loss. This is better than ATRC which looses a 148 byte packet if any one of three ATM 48 byte packets are lost. The AT&T/Zenith proposal cites the I.363 interleave matrix for Reed Solomon correction. This (47+1) x 128 byte matrix would yield 6144 bytes in a correction block. Up to 96 bad bytes, or two bad ATM packets, could be corrected using the 196 error correction bytes out of the 6144 total data bytes.

The ability to correct only 2 out of 128 ATM packets seems insufficiently robust for all but high reliability channels. Testing might demonstrate otherwise, however I would anticipate that an ability to correct 8 or 16 bad packets might be more appropriate for general ATM communications. The ability to correct 16 bad ATM packets out of 128 would require 32 bytes of Reed Solomon FEC instead of the 4 being proposed. The 4 being proposed allows 124 bytes for data. If 32 were used, would the remaining 96 bytes have an efficient mapping into the AT&T/Zenith formats?

In all cases the ability to correct any bad or dropped packets is superior to the ATRC proposal which drops 3 ATM packets, or one ATRC packet, for each ATM packet which is bad or lost.

Since no ATM mapping was proposed by other proponents, the ATM mapping of AT&T/Zenith appears to be the best, since it is superior to other proposal for ATRC.

262 and (2 x 262 - 1=) 523 slice blocks is a strange number per frame, matching the 525 lines of NTSC. Why do such a thing?

Fixed data frame size also seems highly inflexible compared to more flexible ATRC packet formats.

The use of prioritized data is beneficial, as presented in this proposal. However, uniformity of number of levels of priority over various networks and applications will ultimately be required. The two priority level approach of 2VSB and 4VSB may be insufficient to embody the broader priority information contained within the packets. However, any priorities, as in ATRC and AT&T/Zenith, is superior to a single priority as in the other proposed systems.

Header

I didn't understand the header presentation. I believe that the header is buried within the 262 and (2 x 262 - 1=) 523 slice blocks at a specific location internal to the slice. This is not a lowest-level header, and therefore would not serve the function of universal identification. Rather, a decoding system would have to decode the AT&T/Zenith packing structure in order to find and decode the header. Further, the header could only reference one slice of information as proposed, with a single fixed given size.

Apparently there are headers at two levels, one 8 byte header in the first 171 byte slice of a data field, and subsequent potential headers in the video and other data blocks. Only the lowest level is useful as a potential universal header.

The error robustness also did not appear sufficient to support unambiguous interpretation of the header in order for it to perform its function.

This header proposal cannot serve as a universal header

Further work, and potentially an alternate header design, is needed Testing will be required for any serious header proposal.

Augmentation

The augmentation scalability proposal is good as presented. With the except of 59.94 Hz not forming a suitable basis for a family of formats, the progressive scan square pixel format forms a strong basis for augmentation pyramids.

Scalability

The AT&T/Zenith comments on downward scalability were that there is an efficiency penalty. Could this be overcome?

Vector Quantization

Vector quantization of DCT coefficients is likely to improve coding efficiency with respect to other implementations. The MIT switch between three quantizers is a crude version of vector quantization as applied to DCT coefficients. As memory becomes cheaper, larger codebooks will become feasible, allowing further improvements in coding performance.

Leak

The leak enhancements were apparently not tested (regional dynamic leak). The claims for its performance sounded interesting, but certainly this will need to be tested.

Address List

Conditional access methods were described briefly. A proposal for the format of the address list is provided. It should be tested to see if it works before the proposal is accepted. The use and meaning of the fields was not described. It was claimed that 610 authorizations per second could be carried in the proposed format. This seems unacceptably small. Fortunately, this conditional access method does not require picture data, but uses auxiliary data, so the authorizations could be sent well in advance of the event. Provision for handling large numbers of last minute authorizations was described as "fast addressing. It was described that this would provide 5.7 Million subscribers per minute, while providing minimum quality audio/video. Such a facility would require much of the image bandwidth to be applied to address lists for a period of time while a reduced temporal rate or partial screen moving image would be presented.

The quality of such an image should be tested. The number of errors in the 5.7 Million addresses per minute should also be evaluated, together with methods to catch and correct the errors.

MPEG

The ATRC MPEG compression, using bi-directional coding is probably not the best choice. The event last week where AT&T/Zenith and G.I. visited the adhoc MPEG-2 format committee suggests that MPEG-2 may adapt to whatever decision the U.S. makes for terrestrial advanced television.

The fact that Barry Haskell of AT&T indicated that he is chairman of the MPEG-2 entertainment profile committee which is now forming. This tells us that AT&T is serious about participation in MPEG-2.

In my opinion, those who are working on MPEG-2 may attempt to avoid the mistakes made with MPEG-1. However, until MPEG-2 begins to take direction, it is difficult to evaluate level of MPEG-2 compatibility.

MPEG-1 is of such poor quality that compatibility with MPEG-1 coding is of little value.

NTSC Compatibility

The progressive scan 720 lines is easily converted by a simple 2/3 filter to 480 lines of interlace to create a high quality NTSC down conversion. Unlike ATRC and G.I., no motion-vector de-interlacer is required.

720 lines bears a 5/6 relationship with the 576 lines of PAL, making PAL extraction acceptable, athough not as optimal as NTSC extraction.

The M.I.T./A.T.V.A./G.I./C.C.D.C format is identical to the AT&T/Zenith format.

Temporal Rate

The temporal rate of 59.94 Hz, common to all of the proposed systems, is unacceptable for computer display. It is likely to be unacceptable for large bright entertainment screens as well, for the same reasons of broad area flicker.

24 Hz motion picture presentation was not proposed for this system. This is a major disadvantage of this proposal with respect to the other digital proposals, all of which provide for 24 Hz direct image update. Thus the presentation of 24 Hz motion picture film at higher resolution than 59.94 Hz, or other (more acceptable) higher rate imagery on a 70+ Hz display is not provided.

This system should be modified to allow direct transmission of 24 Hz imagery.

Further, 59.94 Hz image update rate should be abandoned in favor of 36 and 72 Hz, or other computer compatible rates.

Partial Image Update

The addendum states digital systems have flexibility to operate only on portions of an image, analog doesn't (a). I agree that analog does not provide such flexibility. However, I am not sure this claimed flexibility is really provided by the digital systems, although I very much hope that it is. As far as I can tell, there are significant differences in the systems with respect to partial image update at a reduced bit rate. (see mention above of 64 x 48 slice tile size being promising for conditional replenishment)

Further, I have seen no proposal for how a transparency channel might be sent and decoded in order to allow a foreground to be animated and composited over a background. This would certainly be one of the preferred methods of partial image update.

MIT

Audio

MIT's four independent audio channels seems most flexible to me. The use of dependent channels is acceptable where a single language is used and where surround sound is desired. However, when multiple languages are desired, multiple independent channels would be needed. A combination of dependent and independent coding might ultimate be most useful. For example, music channels could be coded in a dependent fashion, and mixed with languages contained within independent channels. MIT's independent coding is preferable until a hybrid independent/dependent coding system could be pursued.

Data Format

Does the MIT system have the same G.I. sliding window problem, where panels slide from frame to frame on a 44 or 20 frame cycle?

The odd 525 data line size matches NTSC. Why would this be chosen?

No ATM Mapping was proposed.

No ATM cell loss recovery was proposed.

Headers

Headers and packets are identified by location within fixed sized data blocks. This is a very inflexible use of data transmission formatting. True universal headers have not been defined for this system.

Compression Extensibility

No mechanism was mentioned for dealing with potential compression technique obsolescence.

DCT vs Sub-Band

G.I. claimed that the sub-band without chroma sampling tested more poorly that the DCT with chroma subsampling in an internal test. Could we see test results or tests on prototypes? What was the difference in perceived quality between sub-band and DCT implementations? Wouldn't the sub-band system have been more flexible and scalable? Wouldn't the improved resolution due to no chroma subsampling on the MIT system when showing 24 fps film be noticeable as a significant improvement in picture quality? Was this not shown in the tests which led to the rejection of sub-band? If not, why didn't it look better? Are such differences due primarily to the fact that motion-compensated DCT is more mature, or are these differences fundamental to the coding properties of DCT and sub-band coding?

NTSC Compatibility

The progressive scan 720 lines is easily converted by a simple 2/3 filter to 480 lines of interlace to create a high quality NTSC down conversion. Unlike ATRC and G.I., no motion-vector de-interlacer is required.

720 lines bears a 5/6 relationship with the 576 lines of PAL, making PAL extraction acceptable, athough not as optimal as NTSC extraction

Temporal Rate

The temporal rate of 59.94 Hz, common to all of the proposed systems, is unacceptable for computer display. It is likely to be unacceptable for large bright entertainment screens as well, for the same reasons of broad area flicker.

The 59.94 Hz image update rate should be abandoned in favor of 36 and 72 Hz, or other computer compatible rates.

The M.I.T. system supports direct 24 Hz image update, which could support 72 Hz display. However, this capability was not tested due to restrictions in the ATTC testing process. This capability should be tested.

Scalable temporal rates are not supported.

Motion Vector Block Size

MIT supports both 16 x 16 and 8 x 8 block sizes. Small block sizes such as 8 x 8 or smaller are better for motion vectors, by providing for the future potential of temporal interpolations. Smaller subdivisions, such as 2 x 2 or 4 x 4 would be even more potentially useful. MIT proposed the possibility of using such finer subdivisions within the MIT format. The smaller the region size, the more numerous the motion vectors, and the more accurate the motion vectors, the more feasible that temporal interpolation might become.

This potential for 2 x 2 and 4 x 4 blocks with motion vectors for certain regions of the screen should be fully explored. However, at present, motion vectors used in compression are not considered sufficiently reliable for use in temporal intepolation.

Text and Graphics

The use of overlay planes seems like the most straightforward implementation, as proposed. However, the lack of provision for a specific implementation proposal, and the lack of testing suggests further work. The weaknesses in data formatting of this proposal, as well as provision for near-error-free operation for screen display languages, indicate that design and testing are needed, and very little progress has yet been made.

Still Frames

Still frames not tested.

Picture Data Rate Flexibility

It was claimed that the MIT/ATVA system could operate flexibly at pixel rates up to 100 Million pixels per second. It was also claimed that 24 frame per second film could be transmitted in 6 MHz at double the proposed resolution of 1280 x 720, being 2560 x 1440. The ability to provide increased resolution at 24 Hz should be tested. This capability could be extremely valuable in the presentation of motion pictures, which will be one of the major uses of advanced television.

Augmentation

Augmentation using additional data was mentioned, but no concrete proposal was presented. Higher quality signals at higher data rates were proposed, but have not been tested, except by software simulation in the development lab. MIT proposed 180 Mbits/second, using intra frames only for potential studio use. Others proposed 40 to 80 Mbits/second as potentially being suitable. Investigation of actual suitability would be required before the required data rate can be established. My SMPTE study group on compression described numerous issues with respect to use of compression in a studio environment. Such issues suggest thorough testing.

Encryption

The proposal to use stream cyphers may be inconsistent with data packet formatting or block error correction codes such as Reed Solomon. Concatenation of errors is significant for stream cyphers as it is for block codes. Appropriateness of encryption methods, and their effectiveness and impact on picture quality, need to be investigated and tested.

Cost

MIT/ATVA was the only group which answered my question concerning projected receiver costs. The projected costs seem reasonable.

Unbundling

The MIT system was presented as the most modular of the systems. It was described that all of the components were modular in construction and integration. The modular interconnection with other components of similar function has not been tested, however. On the internal interconnects have so far been demonstrated.

ATRC

Interoperability

The ATRC notion of interoperability is absurd. To claim that the ATRC system is interoperable with may incompatible formats doesn't establish a basis for interoperability among advanced television uses.

The ATRC is the worst of all of the systems in the respect that it takes the absurd notion of interoperability as meaning that it can support multiple incompatible formats and uses. Implied is the notion that these uses will never converge. Computer display on advanced television, and advanced television display on computer would never be easy, since these uses would be supported by differing and incompatible formats. Apparently these problems of interoperability are not important to the ATRC. The choice of an interlaced non-square pixel format to test at the ATTC is an indication that interoperability has not been a high priority at the ATRC.

If the interlaced proposals such as NHK and G.I. were adopted, they would probably be abandoned before very long. If the progressive scanned MIT or AT&T/Zenith systems are chosen, then interoperability is aided, with the exception of the use of the unacceptable 59.94 Hz rate.

However, if the ATRC system were to be adopted, the inevitable conflict between interlace and non interlace, square and non square pixels, would become solidified due to FCC adoption. Thus, this battle would likely rage like a growing problem for many years, causing ever expanding interoperability conflicts. Thus, the highest long term interoperability cost of all of the systems may result from the ATRC system. Because it straddles the border between interoperable practices and non-interoperable practices, it would "let the market decide". This philosophy would cause incompatible, non-interoperable camps to spring up for uses of transmission and distribution media. The ultimate result would be similar to the economic inefficiencies caused by the Beta and VHS war, where ultimately the owners of Beta tape equipment found themselves without new products and services. Similarly, using the ATRC approach, ultimately the interlaced devices and channels may be orphaned. In the mean time, the two worlds of computer and entertainment uses would be incompatible without conversion. The conversion would be costly and would reduce quality. That is primarily the antithesis of the goal of interoperability.

Thus, the ATRC notion of interoperability is precisely the antithesis of true interoperability.

The <u>ATRC system should be rejected as not providing interoperability</u>.

Packet Structure

Even though this system should be rejected in its entirety, it is worth noting some of the more interesting aspects of the design, for possible inclusion in a more acceptable system.

The packet system is easy to understand and well defined. The general notion of a packetized format is probably good. The packet format has potential flaws, however, as proposed by ATRC.

Packet prioritization is good as a notion. The implementation of only two priorities for transmission does not take full advantage of more detailed prioritization available within data. Two priorities are not sufficient for optimized use on congested ATM networks or other applications. Higher numbers of priority levels through transmission needs to be tested.

Droping one ATM packet drops a full Sarnoff 148 byte packet (three ATM packets)
Reed solomon block code used by AT&T/Zenith is much better mapping onto ATM if dropped or bad packets are anticipated. AT&T/Zenith can correct such errors.

Header

The header is at the wrong layer (must decode entire ATRC packet and error correction structure to find header, alternate data formats may not be optimal inside of ATRC packet structure)

The header, as proposed, is not sufficiently robust from errors in order to serve its function. Further, the packets to which the header refers may be independently flawed or dropped. Packets with a flawed header packet or headers pointing to flawed or dropped subsequent packets will likely cause significant data loss. Such issues need to be tested.

The notion that there are headers at two layers is incompatibile with the concept of a Universal Header, as described by the SMPTE task force. The only header which is universal is the outermost layer. Inner layers are not universal headers, but are headers internal only to the specific format proposal.

The notion that receivers should be able to ignore unrecognized headers to provide extensibility is corrrect. However, this is necessary but not sufficient, in order to provide extensibility in any area.

Compression Extensibility

It is claimed that the header/descriptor mechanism in the ATRC proposal could support future compression enhancement. However, the way this might be done is not discussed. This crucial issue needs to be defined and tested, since the likelihood that the ATRC MPEG++ compression format will be obsolete at some point in the future, perhaps as early as five to ten years. Using the header will certainly be required, but how it is used to provide extensibility to new compression algorithms, while remaining on the air with existing receivers, was not described.

Partial Screen Motion

"Partial Screen Motion" is referenced in the addendum. This sounds potentially useful. However, it needs more exploration and testing. It is unclear whether the format as proposed will work or not.

One example is during address list distribution for pay-per view. Pay per view, viewer groupings, video on demand, and electronic billboards are all described, but no mechanisms for accomplishing these features were presented. Such mechanisms should be tested prior to acceptance of the premise that these features can be provided. Encryption, address lists, reduced quality pictures or partial screen update, and other features would be required. No proposal for an implementation of such features was made.

Problems with MPEG

Picture frame groupings are a significant barrier to many uses, and such groupings are fundamental to the ATRC implementation of MPEG. MPEG apparently allows coding without bidirection interpolation, which would be more interoperable and less costly to implement. The bidirectional interpolation may improve the picture quality a small amount. However, the cost in implementation and loss of interoperability may not justify this improvement.

Increased memory is required for bi-direction MPEG-style coding, as well as a more complex decoder.

MPEG style decoding is much less flexible with temporal rates than other coding systems. This is due to groups of frames, such as the ATRC nine frame groupings, not allowing easy implementation of alternate rates. Multiple simultaneous display rates decoded from a common signal does not seem feasible within the ATRC MPEG implementation structure, whereas it might be possible within some of the other compression structures proposed.

It should also be noted that General Instruments and AT&T/Zenith made a presentation to the ad-hoc meeting of MPEG-2 format committee last week (which was the week after our interoperability review). They are proposing that MPEG-2 be developed to be compatible with the HDTV advanced television format proposals and the NTSC encodings based upon these HDTV proposals.

If MPEG-2, which is not yet defined, decides to be compatible with the advanced television system which we ultimately choose, and if a system is chosen other than the ATRC system, then the ATRC system would not be MPEG-2 compatible. Thus, the claimed advantages due to MPEG compatibility of the ATRC system may be unfounded since MPEG-2 may ultimately adopt a system more similar to G.I. or AT&T/Zenith's proposals.

It should also be noted that MPEG-1 is deemed by many to be poor in quality. It was developed and optimized for the low 1.4 Mbit/sec rate of CD-ROM (audio), and it only provides 320 pixels by 240 lines in the common implementation. It is progressive scan and 30 or 24 Hz. Even at this low resolution and modest frame rate, the artifacts are sufficient to cause many people to question whether consumers or professionals will pay for such presentations.

Thus, the concerns over compatibility with MPEG-1 may be unfounded, since MPEG-1 may never become popular due to very poor quality of encoding and presentation.

Issues with respect to compatibility of MPEG-2 are questionable since MPEG-2 is not yet defined, and may ultimately differ substantially from the ATRC MPEG++ implementation. The current MPEG-2 consideration of a tool-box approach may make MPEG-2 more of a syntax than a specific compression format. These questions make MPEG-2 interoperability claims difficult to evaluate.

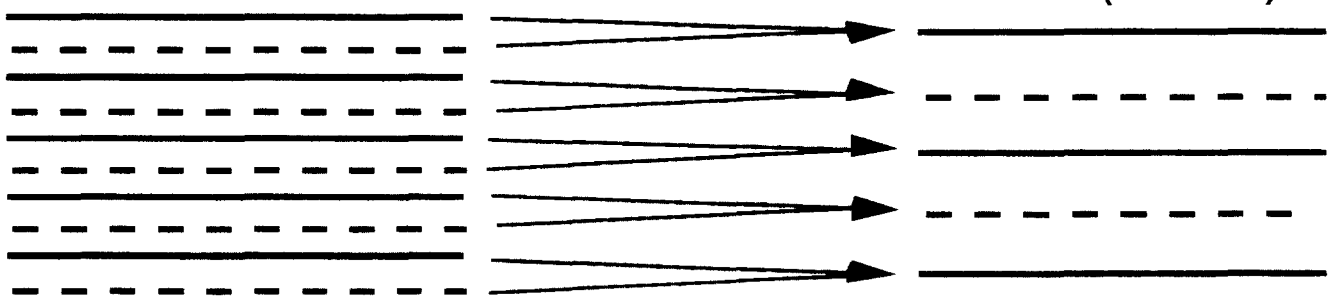
Deflection Rate

The use of 2 times horizontal deflection is described as an advantage. In what context is this an advantage? The only advantages which I can see for 960 lines interlaced are that existing NTSC 480 line interlaced formats can be line doubled using an inaccurate method to generate 960 lines. A correct expansion from 480 lines to 960 lines is complex and involves error-prone motion vector de-interlacing prior to expansion. The quality of upconverted NTSC on advanced television at 960 lines is questionable to begin with.

When attempting to downconvert from 960 lines interlaced to 480 lines interlaced for NTSC, the presence of interlace in the source signal leads to missing information during the down conversion as follows:

Down Conversion from 960 Lines Interlaced To 480 Lines Interlaced

960 Lines (Interlaced) **480 Lines (Interlaced)**



(Odd Lines Solid Even Lines Dashed) **(Odd Lines Solid Even Lines Dashed)**

Problem: The solid lines happen at the same time, and the dashed lines happen at the same time. However, each line in the 480 Line NTSC requires one dashed and one solid line. Therefore, needed information is missing!

Thus, the conversion from 960 lines interlaced to 480 lines interlaced for NTSC becomes very difficult. A motion-vector deinterlacer is required to track all objects in the scene to re-create an estimation of the contents of the missing lines of information before down conversion. Such standards conversion is expensive and error-prone to temporal aliasing due to normal temporal undersampling issues.

The claims that ATRC or G.I. have an advantage using twice the horizontal rate of NTSC vs the three times used in the MIT and AT&T/Zenith proposals seems to be a minor point. The display on my desk for my Apple Macintosh Quadra 700 on which I am typing these comments is 21" display made by Ikegami which extends from 50 to 80 kHz in horizontal rate. The cost of this display was around $2,000. It would have cost me an extra $200 to extend the range downward to 35 kHz so that the display would have gone from 35 to 80 kHz. Such a display is useless at rates below 35 kHz

NTSC operates at 15 kHz, so double this rate would be 30 kHz, which is marginal in the computer context. A 45 kHz rate is more clearly similar to computer uses. However, the only cost penalty apparent was from having a wider multiscan range. Many computers, and probably many future advanced television displays, will use a single frequency. I remain unconvinced that the horizontal deflection repeat rate is a major issue in receiver cost, given that many computer displays exceed 70 kHz and that the cost of these displays does not differ substantially from displays which have rates near 40 kHz. The main cost differences in computer displays occur due to size. The 13" displays cost about 75% of the cost of 17" displays. The 21" displays cost about 60% more than the 17" displays. Size of the screen for CRT's will be the major cost factor, if the analogy with computer display pricing is valid. For projection high definition display systems, it is not clear that the speed of deflection has a significant affect on the cost Certainly for liquid crystal projection, where there is no deflection, deflection is not a factor whatsoever

Use of 1440 Horizontal

1440 is based upon CCIR 601/D1, the first digital video format. CCIR 601/D1 is 720 pixels per scan line for both PAL and NTSC. Unfortunately, 720 pixels does not yield square pixels for either PAL or NTSC. The choice of CCIR 601 was a political compromise between NTSC square pixels at 640 and PAL square pixels at 768. In retrospect, we should have chosen 768, since that would have provided square pixel PAL, but would have yielded 4 times color subcarrier (4fsc) encoding of composite NTSC. As it is, the D2 NTSC digital composite video format must use 4fsc at 768 pixels, and must be converted with a complex line filter from CCIR 601/D1 at 720 pixels, with a resulting degradation in quality.

Although CCIR 601/D1 is popular, its basis in non square pixels makes it unsuitable as a basis for square pixel advanced television which is required for interoperability. I was the first major user of CCIR 601/D1 technology in the United States, producing 1988 election graphics for CBS. The lack of square pixels caused a significant inefficiency in wasted computation due to requirements to have both nonsquare CCIR 601/D1 pixels and square pixels for the computer graphics, doubling the rendering computations.

Thus, CCIR 601/D1 should not be used as a basis for advanced television format parameters, and therefore 1440 is not a good choice.

In addition, even if 720 were a legitimate basis, 1440 is double, not double widened to 16:9 advanced television aspect ratio. Such widening by 4/3 would yield 1920, which is also not suitable for advanced television, due to its basis on 720.

It is also useful to note that 960 vertical lines is not divisible by 9 when building the 16 by 9 aspect ratio. Therefore, the square pixel version of 960 lines comes to 1706.666. This makes 960 lines unsuitable for 16 by 9 aspect ratio formats. This seems to be the reason why 810 vertical lines with 1440 horizontal pixels is proposed as the square pixel format. However, the ratio of 960 to 810 is the complex fraction1.185185... or 32/27. Thus, the numbers 960 and 810 are almost completely non-interoperable. 810 bears a similar disparate relationship to the 480 lines of NTSC, being 16/27 or .592592...

The future production standard mentioned in the ATRC addendum suffers from the same foolish notion as basic notion that supporting multiple incompatible formats promotes interoperability. Progressive scan non-square pixel 1050 progressive at 59.94 Hz is not useful because of both 59.94 Hz and the non-square pixels. This is proposed as a future standard, with 29.97 Hz progressive and 59.94 Hz interlace for now, neither of which is interoperable with computer and other needs The Atrc proposal is not in touch with the realities of computer and other interoperability

The receiver capabilities addendum postulates low cost interlaced receivers, and high end 72 hz progressive scan displays running film. It was not observed, however, that the proposed signal could not be received on both devices simultaneously. Either there is a 24 hz progressive scan, for 72 hz playback, or there is a 960 line interlaced scan being transmitted, which must be played at 960 lines interlaced. This 960 line interlaced signal cannot be acceptably played back on a 70+ Hz progressive scan display, as required for computer display interoperability. The ATRC proposals are thus are not interoperable with all receivers, but rather with a single class of signal at any one time. This is likely unacceptable to any broadcaster or service provider, since they would not be broadcasting to all receivers, but only those of a particular class, at any given time.

VCR

The VCR pause comments in the ATRC addendum do not seem credible to me. The MPEG structure will likely cause pausing to only be available every 9 frames at "intra" frames. Such pausing would not work well to the individual frame.

The VCR packet size proposed in the ATRC addendum doesn't seem credible either, in that multiple reed solomon blocks, similar to those used in AT&T/zenith are likely. No mapping into such larger reed solomon correcting blocks is discussed. The ATRC packets do not have appropriate correction codes for VCR implementations due to small packets. Such small packets without packet groupings are likely to be destroyed by normal burst errors on VCR devices.

G.I.

Sliding window problem, where panels slide from frame to frame on a 44 or 20 frame cycle?

does this problem exist in mit system as well?

Large Macro Blocks for motion vectors, coarse motion vector independence, reduced fleixibility

Motion vectors not resolved to sub-pixel accuracy

very heavy chroma subsampling (4 to 1 horizontally). could cause some severe loss of sharpness for red or blue objects against black, or yellor or cyan objects against white.

525 is odd number of cells per frame, corresponding to NTSC scanlines, why do such a thing?

Concerns about cost in addendum seem misplaced (1). The whole point of scalability is to have reduced cost decoders. The current gi proposal is a single format system. Due to interlace and non square pixels, it is a particularly difficult format to use for lower cost lower resolution receivers. Also, due to interlace and non-square pixels, the format is poor for higher quality future systems as well. It does not work well in an augmentation pyramid.

Proposals that the 960 interlaced format be converted to square pixels and progressive scan (4), with resulting temporal and spatial resolution quality loss, are very unattractive.

The suggestion that 1080 lines would be more popular that 720 for progressive scan is based on the 1920 x 1080 common image format proposal. In my well aired opinion, 2048 x 1152 is a much better choice. 1920 is viewed by many as an "odd" number for a digital system, as is 1440 or 1408. Numbers such as 2048, 1280, 1024, and 640 are much more easily recognized for horizontal pixel resolutions as being digital computer compatible numbers.

The proposal that 24 fps progressive imagery could be displayed at 72 Hz on a computer display is a good one. However, the non square pixels force a conversion and loss of quality. Such a proposal would be much better if the GI format were square pixels.

With respect to digital busses, the EIA CE bus was mentioned (5). Is this bus sufficient for real time at the required bandwidths? If so, shouldn't its use be tested? What about IEEE P1394 (Serialbus) and P1596 (SCI) as interconnects?

I don't understand the description of scalable data rates as being linked with non-real time and decoupled clocks (6). Isn't the reason for scalable data rates to allow simultaneous synchronized lower cost reception of the same signal?

Still frames are described as possible (7). Needs to be tested. How many per second, what quality. How can an interlaced system be used for still frames, or must it run in progressive mode for still frames. Thus, no still frames from sports coverage. What about sliding window issue?

The transwitch proposing discussion (10) for cable addressing proposes using a reduced data rate for images during the addressing time. Still frame is proposed, but couldn't a partial screen update or slow moving image be supported? Should be tested to see if this works within this system.

I chaired a compression study group in SMPTE concerned with studio and contribution quality uses of compression. Our final report is pending approval in SMPTE. When it is released, it would be useful if each proponent received a copy. If each proponent desires a preliminary copy, such can be provided with Ken Davies approval as SMPTE engineering vice president. The study indicates that studio and contribution quality uses of compression can involve numerous complex issues. These issues are outlined in the report. At the 30 to 45 Mbit/second rates discussed (11), numerous potential problems can cause severe artifacts for production and contribution uses. Care should be taken to demonstrate any proposed production or contribution level of compression with typical production and contribution post-production uses such as titling, re-framing, contrast or brightness adjust, color adjust, matteing, standards converting transcodings, etc. In the absence of testing, it is difficult for me to verify that G.I. claims that 30 to 45 Mbits/second using the G.I. compression system would be suitable for such uses as claimed.

Surround Sound (12) can be provided for more than 4 dependent channels. Two independent channels can be supported, which would presumably consist of two languages, each in monaural. We have not had any discussion of methods to support more foreign language channels. Two stereo channels, separation of language of dialog and music, or more foreign languages could be quite desirable. The MIT proposals of 4 independent channels seems more flexible.

The use of an interlaced, non-square pixel format as a basis for augmentation using residuals from feedback loops is far inferior to progressive scan and square pixels. The augmentation channel containing the resolution or temporal rate enhancing residual is much more efficient when derived for progressive scan. Square pixels provides anisotropy and compatibility with industry (correct) practices. The augmentation residual is much more highly compressible due to superior coherence and reduced residual values when a progressive scan source is used instead of an interlaced format. Thus, I do not feel that the G.I. interlaced format could be used effectively as claimed as a basis for augmentation to higher quality levels.

The use of panels (15) apparently creates a sliding window problem, since the panels are not phase locked to the frames in position, but rather precess across the image. This makes the manipulation of packets, slices, macroblocks, or other structures, when attempting to control image update on a regional basis, quite difficult. Data structures should be statically linked to image regions so that image regions can be easily manipulated independently of one another.

The relationship of macroblocks, 525 data lines, panels, and other G.I. structures are unclear to me, since panels apparently precess. The use of headers at the start of each frame of 525 data lines must therefore have a precessing relationship to panels. This is a complex structure to try and decode in order to find and interpret headers.

Video Packet and Transwitch Format Layers Discussion

The word Transwitch is really just a directed multiplexor (mux) function to separate audio, video dct components, motion vectors, etc. Sync is needed with packets, as proposed. However, the packetization used in the G.I. system is not computer-like, but is much more of an internal format type of data stream. The ATRC system packets were more explicitly designed for external manipulation and flexibility of use.

The number of bits in the headers are very small. Further, these headers are only headers to extremely fixed sizes of packets. Thus, the header proposal is probably substantially under powered to serve as part of a universal header system. The need to decode the unusual 525 line frame and packet structure of G.I. in order to find and interpret headers makes the headers specific to this format, and probably not adaptable to other uses. No provision in the G.I. system for foreign headers with foreign data sizes of packets is provided.

Reference is made to the transwitch layer being used as part of the buffer fullness feedback loop which enables constant rate coding. However, it is also claimed that this same system could be used for packet networks (like perhaps ATM?) to pack different sized cells for different network uses or conditions. It is not clear how this would work, or whether such repacketization using this system would have desirable characteristics.

The two layers, being the frame and the data line, have headers. This fixed two level hierarchy of fixed-sized packets seems highly inflexible. Are there provisions for more data lines per frame than 525? Can there be less than 525 lines where some lines are used for other purposes, not presently defined?

It appears that much of the header information is used to point to panels, macroblocks, headers, frames, and lines. The reason is in case packets are lost, so that orientation within the data can be reconstructed. Although this may be necessary, it does not seem warranted to claim that these structures serve more general purposes at the same time.

NHK

The analog nature of the NHK system, together with interlace and non-square pixels, lack of a header, etc. lead me to reject this proposal as entirely unsuitable for interoperability.

Unbundling

(also called "mix and match", or "combination system")

Modem

The transmission of data over formerly analog communications channels is essential the function of a modem (modulator/demodulator). The modems in these advanced television proposals appear to work acceptably, although the testing results have not been shown. There are variations in the modulation schemes which may make certain of them more suitable for various puposes. For example, 32qam may be acceptable for cable or satellite. However, the ATRC and Zenith co-channel avoidance approaches appear on paper to be superior for terrestrial broadcast.

AT&T/Zenith indicated that their use of variable bit rate in transmission ties their transmission system to the compression system, thus they are not separable. The audio is separable (dolby), but it is unclear regarding the separability of other aspects of the system.

However, it may be possible to separate the modem technology from the formats, temporal rates, data packings, etc. If such is possible, we should select the modem technology which provides the most reliable and highest bandwidth data delivery.

It is unclear how modular the ATRC modem might be, but if it were separable, it might offer one of the better choices.

MIT claims that their system is very modular.

G.I. provided no discussion of this issue. However, MIT sponsorship implies some modularity for the MIT incarnation of G.I. built portions of MIT/GI/ATVA/CCDC system.

Audio

MIT audio has 4 independent channels. Even though it uses more bits per second, the independent channels allow more flexibility. For example, 4 simultaneous different audio languages could be sent using the MIT audio system.

Square Pixels

Only Square pixel scanning parameters should be considered.

Progressive Scan

No interlaced formats should be allowed. Interlace is the largest single barrier to national and international interoperability.

Scanning Parameters

Scanning parameters, for those who addressed them, seem to be quite flexible (see my recent paper entitled "A Scalable Family Of Formats For U.S. Advanced Television"). Jae Lim made particularly flexible claims in this area. The ATRC claimed to support several quite different (and incompatible) scanning parameter formats. Unfortunately, they did not continue their support for the most computer friendly of their proposed formats at 1820 x 1024.

Although 1280 x 720 is not the best choice of resolution scanning parameters, it may be acceptable. No formats based upon 1440 pixels horizontal seem to be acceptable

We should use the 24 Hz native encoding offered by MIT, G.I., and ATRC. In particular, MIT indicated that they could double their resolution to 2560 x 1440 when showing 24 Hz film. Although I do not favor 2560 x 1440, I would like to see 2048 x 1152 suppored at 24 Hz.

I would also like to see Photo-CD and 4096 x 2304 supported for non-real time still frame applications, as well as other aspect ratios.

Temporal Rate

The strong bias toward 59.94 Hz by the proponents and by the ATTC leads to a lack of interoperability with computer CRT displays, which require a refresh rate in excess of 70 Hz.

Some of the proponents indicated that they could adjust their scanning parameters within the 20% that would be required to go to 72 Hz. Such an adjustment may become critical for computer interoperability.

Temporal rate scalability or flexibility should be based upon integral multples of 1/72 of a second. Thus 12, 14.4, 18, 24, 36, and 72 Hz could all be supported as temporal rates.

Scalability

None of the systems as proposed is scalable. Building reduced resolution subsets of the advanced television formats would be very worthwhile in reducing receiver cost. This would be most important during initial introduction, when receiver cost might be prohibitive without reduced-cost models. Even though a reduced-cost scalable subset of advanced television is being discussed, such a subset might have substantially superior quality to NTSC composite television. Potential improvements include component signal, digital encoding, wide screen aspect ratio, better audio, progressive scanning, square pixels for computer compatibility, and a higher temporal rate than NTSC

Augmentation

I like augmentation proposals of MIT and AT&T/Zenith. We should adapt such an augmentation proposal to provide extensibility in resolution. Some form of augmentation might also be useful for providing extensibility to the temporal rate.

ATM Mapping and Cell Correction

I like AT&T reed solomon blocks which could correct dropped ATM packets. However, I suggest that more bytes be allocated such that 8 or 16 dropped ATM packets out of 128 could be corrected instead of the 2 corrected ATM packets as proposed.

Slice Sizes

The slice sizes of most of the systems, based upon 525 data blocks, seem odd. I see no relationship between the 525 scanlines of NTSC and the data blocks. I suggest that new blocking and slice formats be devised with a broader interoperability context in mind.

Packet Formats

The ATRC packet formats seem more flexible and easier to understand. However, they will probably need to be adjusted.

Conditional Replenishment Cell Size

I like the AT&T/Zenith slice cell 64 x 48 pixel independent coding units. Such screen tilings represent a good step toward conditional replenishment or partial screen update. However, the actual use of tiles for such purposes should be thoroughly tested.

Header

None of the header proposals appears to be able to provide the functionality of a universal identifier. Additional design work will likely be required before an acceptable universal header descriptor, as described by the SMPTE task force, can be developed for use with advanced television.

Compression Efficiency

None of the proposals has a method to allow for extensibility to future superior compression algorithms. Such a mechanism should be developed and built into the format of the system ultimately selected. Programmable decoders, skipped data areas, shrinking compressed picture areas, and other techniques might be useful mechanisms.

Questions From Gary Demos
FCC WP4 Review Board Member

Generic Questions For The All-Digital Proponents

*** - Key Questions Marked With Three Stars - ***

Computer Screen Text And Graphics Display

*** 1. *How would you propose to handle non-band-limited image data? Such data is typically presented on computer screens as text, window borders, and graphics. Such data is typically raster-aligned and is usually created from a source format such as run-length coding, Adobe Postscript (tm), Apple Macintosh quickdraw (tm), X-Windows (for Unix), or Microsoft Windows (tm) (for PC DOS). Would you use overlay planes in the receiving device? Would you put this image data through your compression algorithm?*

2. *Are the data areas in your system sufficiently robust, or could they be augmented with further error correction such they might contain graphics screen data as described above (in 1)? Postscript and other screen or printer formatting data types are extremely intolerant to errors. How could errors during screen display be handled if they were to occur?*

3. *What mechanism should be used to decide what screen presentation language formats should be supported by advanced television systems? Should the FCC decide using one or more non-proprietary data representations, or should proprietary and non-proprietary display data representations be supported? What complexity would be associated with supporting more than one such format for interpretation to the display?*

4. *Would you expect to see numerous fonts supported? What would be the additional receiver display system complexity associated with a large number of fonts vs a few simple fonts?*

*** 5. *Jim Clark, the founder and Chairman of the Board of Silicon Graphics, says that we should plan for 3-D computer graphics in the receiving display. How do you think such 3-D graphics would be integrated with the advanced television picture? How would you expect to support 3-D graphics, and what data formats would you expect to use? Is the data area appropriate and sufficient for this purpose?*

6. *If 3-D graphics were to be used for a backgrounds, and advanced television were to be used for foregrounds, a transparency (or alpha) channel would be required in order to provide anti-aliased matte edges. How would you anticipate sending a partial image together with a transparency channel for such a 3-D and advanced television hybrid moving image?*

Unbundling Of System Components

*** 1. *If your system were to be judged the best in some aspects but not in others, could these portions of your system be utilized with portions of other proposed systems? For example, one transmission and modulation system might be the best at reliably delivering digital data. A different system might have the best error correction. Yet another system might have the best compression/decompression system. Yet another might have the best scanning parameter formats. Could you identify which portions of your system might be "unbundled" and which portions must work together due to tight integration requirements?*

*** 2. *Is every proponent's system somewhat independent of scanning parameters, within limits? What do you estimate the range of such flexibility in scanning parameters might be?*

Resolution Scalability

*** 1. *Could your system be expanded to include resolution scalability in the compressed format?*

*** 2. *If so, could you accommodate a lower resolution image at somewhere near 864 x 486 for 16 x 9 aspect ratio wide screen and 648 x 486 for the commonly used 4 x 3 aspect ratio? Could you support decoding these reduced resolution formats without requiring memories and decoding rates associated with the higher advanced television resolutions? Do you feel that some scalable structure such as this would have value in advanced television for the United States?*

3. *Some proposed systems have a higher priority portion of the transmission (ATRC and AT&T/Zenith). Could the reduced resolution format which is described above (in 2) be sent primarily in higher priority?*

4. *Cable Labs is seeking proposals for four compressed NTSC signals in one 6 MHz channel. In your opinion, could or should the lower resolution image in a scalable resolution system (as mentioned above in 2) be the same format of compressed signal, or a similar compatible format, as is used to compress multiple channels into 6 MHz?*

*** 5. *What is your estimate of the cost of a receiver for each year during the next decade for your full advanced television format? What is your estimate of the cost of a receiver for a reduced resolution format as mentioned above (in 2)?*

6. *If a second 6 MHz channel were to be allocated for augmentation of the advanced television picture, could your system be expanded to offer higher resolution beyond the first 6 MHz advanced television format?*

7. *How important do you think square pixels and progressive scanning may be in creating resolution scalability?*

8. *How could your system be used to send a color still image to a color printer? Could the data area be used in conjunction with the main picture stream to provide this capability?*

Temporal Rate

*** 1. *Can your proposed system provide a 24 frame per second image stream from motion pictures, such that a 72 Hz refresh display could be used?*

*** 2. *Computer cathode ray tube displays require a refresh rate which exceeds 70 Hz. How can your system be used or modified to allow presentation of advanced television on computer displays in the home or office? Would there be motion artifacts in such a presentation, and if so, how problematic is their appearance?*

*** 3. *If both 59.94 and 60.0 Hz are found to be unworkable for these reasons, could your system be adapted to 72 or 75 Hz? How big of a modification would be required and what would be the expected performance?*

*** 4. *If temporal rate compatibility with computer cathode ray tube displays is deemed to be critical, can these temporal rate issues be tested with your system? How much time and effort might such testing take?*

Temporal Rate Scalability

*** 1. *If a reduced resolution format such as 864 x 486 and 648 x 486 were to be imbedded within a two or more level scalable system, what temporal parameters should these resolutions have? Would 24 frames per second (progressive scan), like film, be most universal, or would another temporal rate be best? Would a progressive scan version of these formats be more optimal than an interlaced NTSC-like format, by allowing both computer display as well as NTSC-interlaced display to be derived from the progressive scanned format?*

2. *Could the use of motion vectors and compressed corrections, which is common to all proposals, be used to create a hierarchy of temporal decoding rates? For example, could 24, 36, and 72 Hz image update rates all be decoded directly from the same scalably compressed format?*

3. *Can your system be modified or adapted to update different regions of the image at different temporal rates? If your system already does this in a hidden fashion, is it possible to provide this "conditional replenishment" update capability more explicitly to more fully optimize image presentation?*

*** 4. *Some of the proposals can send 24 frame per second film images at the native 24 frame per second rate. This rate is substantially slower than the 29.97 or 59.94 frame rates that have been tested. How can your system be adapted to provide higher resolution at 24 frames per second than for higher frame rates? In your opinion, would such higher resolution at 24 frames per second provide an enhanced viewing* *experience of movies over using the same resolution for 24 frame per second material as for the higher rates?*

Channel Capacity Scalability

*** 1. *During the next ten to twenty years, fiber communication will provide bandwidth to homes and offices in the hundreds or thousands of megabits. The proposed advanced television formats use approximately 20 Mbits/second. How should capacity of hundreds or thousands of megabits be best used? Can your system easily scale to use 40 or 80 Mbits/second? What format parameters would you improve? Examples include wider gamut colorimetry and dynamic range (more bits per pixel), wide screen aspect ratios, higher resolution, more sound channels, higher temporal rate, stereoscopic display, multiple screens, parallel information channels, etc.*

2. *If greatly increased digital capacity were to be available in the future, could you adjust your system to use variable data rate in order to provide more constant image quality?*

Channel Interoperability

1. *In the ATRC proposal, 148 byte packets are used with a 128 and 120 byte payload. How would such packet schemes be related to such protocols as ATM which uses 53 byte packets with a 48 byte payload? How would packet priority be used with ATM or other such systems? How would packet reordering or dropping be handled? For those proposals which do not use packets, how would you propose to send data over a network such as ATM?*

Data Encryption

*** 1. *It may be desirable to encrypt the advanced television data in order to protect the image and sound from unauthorized viewing. Although the advanced television proposed systems are being tested in the presence of data errors, they are possibly not being tested in an encrypted form. What encryption algorithm for your system's data do you favor? How sensitive is such an algorithm to errors in the data? How would data errors affect the picture quality since the data errors occur in the encrypted data stream? How would networks such as ATM networks with potential packet reordering or dropping affect encrypted data?*

Interactive Two-Way Communication

*** 1. *Current television delivery via terrestrial broadcast, satellite, and cable is predominantly one-way broadcast. When fiber systems come into existence, two way interactive communication will become feasible. How would you best make use of this two way or interactive communication capability with your advanced television system?*

*** 2. *What is your estimate of the cost of an originating workstation for teleconferencing in your advanced television format over each of the next ten years? If you supported a reduced resolution format such as 864 x 486 or 648 x 486 as part of a scalable system, what would you estimate the cost of an originating workstation using only this resolution for moving images. but using your full advanced television resolution for drawings, fax, and white-board communication?*

Storage Media

*** 1. *How would you propose to format advanced television on video tape and video disk type devices? What would be the likely affect of media errors? How might fast forward and fast reverse be implemented?*

2. *Is it feasible to have scalable quality levels for media such as video tape and video disk? Could useful advanced television be presented at 5, 10, 20. 40, or 80 Mbits/second to provide various cost/quality and play length levels?*

Compression Efficiency Extensibility

*** 1. *Digital image compression technology, upon which all of the digital HDTV proposals are based, is a rapidly advancing field. Technical developments in just the last two years have seen major new developments and improvements in compression quality and efficiency. This trend is likely to continue for many years. How can your use of data compression take into account rapid major advancements in compression techniques? Can you devise a method to extend your system by upgrading to new more efficient compression while not resulting in immediate obsolescence for those receiving displays using the currently proposed compression technique?*

2. *Do you anticipate that decompression chips in receiving displays will be programmable to some degree? How would you take advantage of such programmability?*

Use Of Header/Descriptor

*** 1. *In the proposed advanced television systems, the packet and error protection structure is such that these are placed at the outer most layer. One goal of the header/descriptor is to help identify unknown data streams. For this purpose, it was originally conceived that the header/descriptor would be the outer-most layer. How could your system accommodate the header/descriptor as an outer-most layer? If you intend for the header/descriptor to be an inner layer, how would you propose that it serve its universal identification function for data streams?*

*** 2. *None of the proposed advanced television systems require error free transmission of picture data. Audio data and data within the data area may need to be almost error-free. The header/descriptor must also be interpreted without errors in order to function properly. Redundant transmission, error-correction-interleaving, and a separate transport header are possible mechanisms. How might your system expect to support the error-free header interpretation requirement?*

*** 3. *In some of the proposals, data is grouped into packets which are prioritized. In all proposals, the data contains the separate elements of audio, picture brightness, color, motion vectors, data areas, etc. How would you propose to use the header/descriptor to identify each such data sub-area?*

Questions For Proponents Of Progressive/Square Pixel Formats (MIT/ATVA/G.I. & AT&T/Zenith)

*** 1. *The MIT/ATVA/G.I. and AT&T/Zenith format of 1280 x 720 will use a production format of perhaps 1312 x 738 to allow extra border for image processing. A resolution hierarchy for the lower resolution image of 864 x 486 and 648 x 486 would be based upon the scaling fraction 2/3. Most scalable image resolution hierarchies have been based upon 1/2. Do you feel it is feasible to build a scalable resolution compression hierarchy based upon a 2/3 scaling relationship?*

Questions For Proponents of Digital Interlaced Formats (G.I., ATRC)

*** 1. *How would you present non-band-limited image data on the interlaced display? Would the image presentation be limited to text and graphics in which horizontal features span at least two or four lines? Would this be done by magnifying existing text and graphics by a factor of two?*

*** 2. *If you had to revise your format to have square pixels and be progressively scanned, what format would you favor? Formats mentioned in the ATRC proposal include 1440 x 810 and 1820 x 1024. Would these be preferable to the 1280 x 720 format? At what image update rate would you run 1820 x 1024? How could each of these be best fit into a scalable resolution hierarchy?*

*** 3. *The square pixel format of 1440 x 810 would require a 3/5 scaling ratio to be scalable to 864 x 486 and 648 x 486. Most scalable image resolution hierarchies have been based upon the scaling fraction 1/2. Do you feel that it is feasible to build a scalable resolution compression on a 3/5 scaling ratio?*

*** 4. *There are many who feel that horizontal resolutions such as 4096, 3072, 2048, 1536, 1024, 1280, and 640 are most desirable for digital display systems due to the match between these resolutions and digital chips and circuits. There is therefore some sentiment that television systems based upon the CCIR 601 horizontal resolution of 720, which include 1408, 1440, and 1920, are not appropriate for many industries. Could your system's use of 1440 or 1408 be adjusted to either of the nearby values of 1536 or 1280? Square pixels would yield 1536 x 864 in addition to the more familiar 1280 x 720.*

*** 5. *How difficult would it be to test a square pixel progressively scanned format using your proposed system? Is such testing feasible if it were to become a critical issue?*